Received: 17 January 2019 | Accepted: 17 April 2019
DOI: 10.1111/jfb.13987

REGULAR PAPER

JOURNAL OF FISH BIOLOGY fsbi

DNA barcoding uncovers extensive cryptic diversity in the African long-fin tetra *Bryconalestes longipinnis* (Alestidae: Characiformes)

Jairo Arroyave[1,2] | Christopher M. Martinez[3] | Melanie L. J. Stiassny[2]

[1]Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad de México, México

[2]Department of Ichthyology, American Museum of Natural History, New York, New York, USA

[3]Department of Evolution and Ecology, University of California, Davis, California, USA

Correspondence
Jairo Arroyave, Instituto de Biología, Universidad Nacional Autónoma de México, Circuito Zona Deportiva 53, Ciudad Universitaria, 04510 Coyoacán, Ciudad de México, México.
Email: jarroyave@ib.unam.mx

Funding information
Financial support for the study was provided by the AMNH Axelrod Research Curatorship to M.L.J.S.

Abstract

To investigate the presence of cryptic diversity in the African longfin-tetra *Bryconalestes longipinnis*, we employed DNA barcoding in a phylogeographic context, as well as geometric morphometrics, documenting for the first time genetic and body shape variation in the species. Analysis of cytochrome oxidase I gene (*coI*) sequence variation exposed extremely high levels of genetic differentiation among samples from across the geographic range of the species (up to 18%), certainly much greater than the traditionally employed *c.* 3% sequence divergence heuristic threshold for conspecifics. Phylogeographic analyses of *coI* data revealed eight clusters/clades that diverge by >4% and up to 18% (*p*-distance), potentially representing cryptic members of a species complex. A clear biogeographic pattern was also uncovered, in which the two main *coI* lineages corresponded geographically with the upper Guinea (UG) and lower Guinea (LG) ichthyofaunal provinces of continental Africa, respectively. Within each of these main lineages, however, no apparent phylogeographic structuring was found. Despite strong genetic differentiation, there is considerable overlap in body shape variation between UG and LG populations. For the most part, morphological variation does not match the strength of the molecular phylogeographic signal. Therefore, the ability to reliably utilise external body shape for regional delimitation remains elusive. Further anatomical investigation appears necessary to establish whether compelling diagnostic morphological features do exist between the divergent lineages of the *B. longipinnis* complex uncovered in this study.

KEYWORDS
Brycinus, *Bryconalestes*, *coI*, geometric morphometrics, mtDNA, phylogeography

1 | INTRODUCTION

The African long-fin tetra *Bryconalestes longipinnis* (Günther, 1864) (Figure 1), is one of over 120 species in the Alestidae, the most diverse of three characiform families endemic to Africa (Arroyave & Stiassny, 2011; Calcagnotto *et al.*, 2005). A relatively small-sized (*c.* 8–12 cm) alestid, *B. longipinnis* is recognised phenetically only by a combination of morphological features including the presence of a fronto-parietal fontanel retained in adults, six teeth in the outer premaxillary row, 5.5 scales in transverse row between lateral line and dorsal fin origin and 8.5–11.5 predorsal scales (Géry & Mahnert, 1977; Paugy, 1986; Paugy & Schaefer, 2007), but no autapomorphic features are currently known to diagnose the species (Zanata & Vari, 2005). Widely distributed throughout coastal drainages of northern West Africa, from Gambia to the Rep. Congo, but absent from the Congo Basin (Figure 1), *B. longipinnis* is generally found in lower

 © 2019 The Fisheries Society of the British Isles

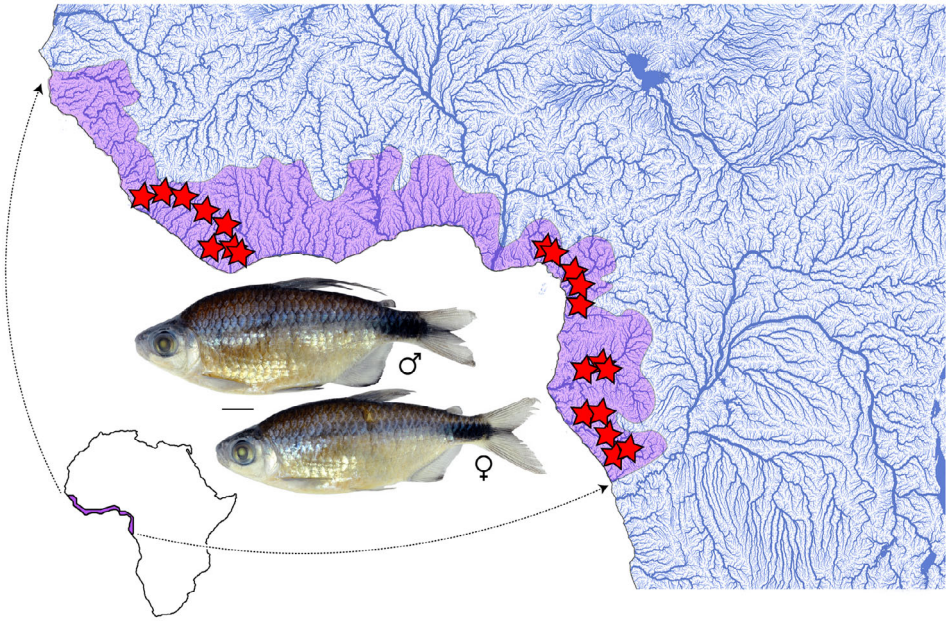

FIGURE 1 Photographs of *Bryconalestes longipinnis* showing sexual dimorphism and map of the species' geographic distribution. (■), Extent of occurrence; (★), Sampling localities. Inset scale bar = 1 cm

reaches of large rivers, including brackish waters of estuarine environments and it is the only *Brycinus*-like species known to penetrate small rivers and streams, although such individuals are generally smaller than those found in large rivers (Lalèyè & Moelants, 2010; Paugy, 1982, 1986; Paugy & Schaefer, 2007). The species is of some economic importance and is harvested both for subsistence consumption and for the aquarium trade (Lalèyè & Moelants, 2010).

Originally described as *Brachyalestes longipinnis* by Günther (1864), but most often referred to under the generic name *Brycinus* Valenciennes 1850, the taxonomic history of *B. longipinnis* is convoluted. The name *Bryconalestes* was proposed by Hoedeman (1951) to include *Brycinus longipinnis longipinnis* and *Brycinus longipinnis chaperi*, in one of the earliest attempts at providing a classification scheme for *Brycinus* and related genera. Subsequent authors investigating the taxonomy of *Brycinus*-like alestids, however, did not recognise *Bryconalestes*, but rather retained the species within *Brycinus* (Géry & Mahnert, 1977; Paugy, 1982, 1986). In what constitutes the sole revisionary study of the group, Paugy (1986) proposed three species assemblages within *Brycinus*, identified as the '*longipinnis*', '*macrolepidotus*' and '*nurse*' groups. These assemblages were delimited phenetically, primarily based on the relative position of the dorsal fin and fronto-parietal fontanel, body size and the number of teeth in the outer premaxillary tooth row. Zanata and Vari's (2005) morphology-based phylogenetic analysis of the Alestidae, however, redefined the limits of *Brycinus* by assigning most species in Paugy's '*longipinnis*' group [*i.e.*, *Brycinus bartoni* (Nichols & La Monte 1953), *Brycinus derhami* (Géry & Mahnert, 1977), *Brycinus intermedius* (Boulenger 1903), *B. longipinnis* and *Brycinus tholloni* (Pellegrin 1901)] to a resurrected *Bryconalestes* (see also Eschmeyer *et al.*, 2018), a taxonomic reassignment supported by phylogenetic hypotheses of alestid intrarelationships based on DNA sequence data (Arroyave & Stiassny, 2011; Calcagnotto *et al.*, 2005; Hubert *et al.*, 2005).

Although considerable morphometric, meristic and pigmentation variation has been recognised among populations of *B. longipinnis*, this was considered to be 'clinal' by Paugy (1982, 1986) and disregarded as evidence for further taxonomic subdivision. While it is possible that *B. longipinnis* is indeed a single, widely distributed, morphologically variable species (Figure 2), recent studies of other such widespread African freshwater fish species, including some of Africa's most iconic fishes [*e.g.*, the electric catfish *Malapterurus electricus* (Gmelin 1789) (Norris, 2002), the African pike characin *Hepsetus* spp. (Decru *et al.*, 2012, 2013, 2015), the African butterflyfish *Pantodon buchholzi* Peters 1876 (Lavoué *et al.*, 2010) and the tigerfishes, *Hydrocynus* spp. (Goodier *et al.*, 2011)] have uncovered considerable actual or potential cryptic diversity. Furthermore, alestid species are one of Africa's notoriously difficult taxonomic groups, characterised by problematic generic and species assignments mostly lacking morphological diagnoses, poorly defined geographic distributions and a taxonomy that is in almost complete disarray (Schaefer, 2007; Toham & Teugels, 1997). Therefore, it seems reasonable to consider the possibility that species-level diversity in *Bryconalestes* may currently be underestimated and that *B. longipinnis* corresponds to a complex of genetically, morphologically and geographically distinct lineages.

To assess phylogeographic structure and possible cryptic diversity in the widespread species *B. longipinnis*, this study employed DNA barcoding, a molecule-based species identification method that, in animals, uses partial fragments of the mitochondrial cytochrome c oxidase I gene (*coI*) as species tags and has proven effective at detecting both described and undescribed fish species (Decru *et al.*, 2016; Hubert *et al.*, 2008; Lara *et al.*, 2010; Lowenstein *et al.*, 2011; Valdez-Moreno *et al.*, 2009; Ward *et al.*, 2005). Specifically, comparative DNA sequence data in the form of *coI* barcodes are used here to test the hypothesis that populations of *B. longipinnis* across its geographic range are indeed conspecific, in which case *coI* haplotypes should not be expected to differ much more than the heuristic standard threshold of *c.* 3% sequence divergence (Avise, 2000; Hebert *et al.*, 2003, 2004; Pereira *et al.*, 2013;

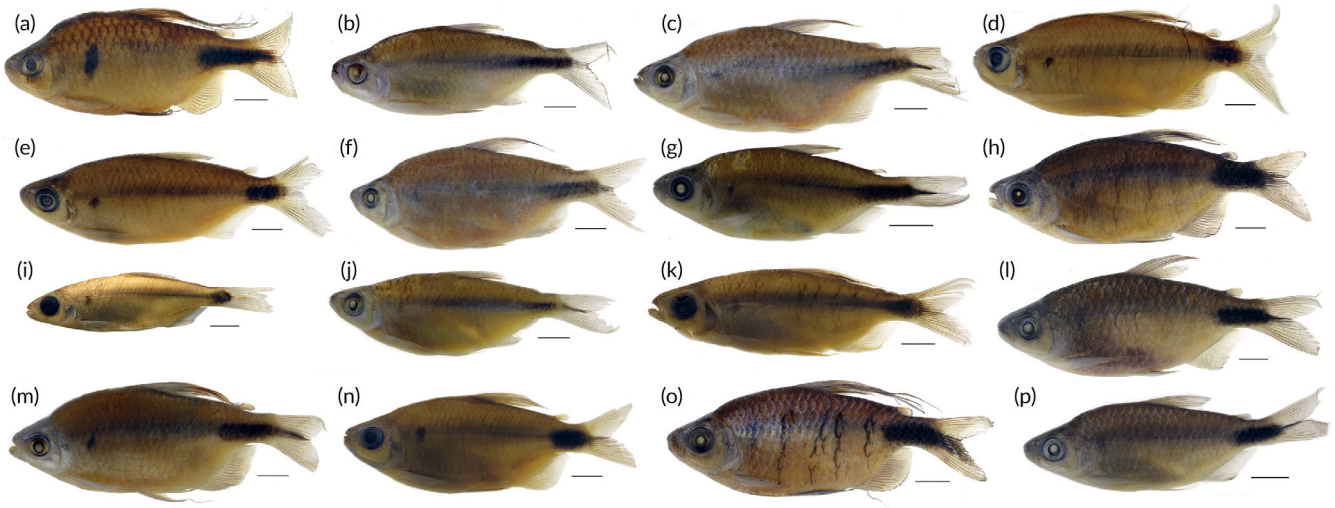

FIGURE 2 Morphological variation in *Bryconalestes longipinnis* across drainages and groups (cluster–clades) representing potential cryptic species: (a) American Museum of Natural History (AMNH) 263263, male, group I; (b) AMNH 259205, male, group II; (c) AMNH 260951, male, group II; (d) AMNH 263201, male, group III; (e) AMNH 263184, female, group III; (f) AMNH 257800, male, group IV; (g) AMNH 254050, male, group IV; (h) AMNH 263041, male, group VI; (i) OS 19395, female, group VII; (j) AMNH 253875, female, group VIII; (k) OS 19393, male, group VIII; (l) AMNH 258328, female, group V; (m) AMNH 249831, male, group V; (n) OS 19397, male, group V; (o) AMNH 258139, male, group V; (p) AMNH 258406, female, group V. Scale bar = 1 cm

Song *et al.*, 2008; Ward, 2009) (exceptions to this heuristic threshold, however, have been reported in African freshwater fishes such as the Malawi haplochromine Cichlidae [Joyce *et al.*, 2011; van der Walt et al., 2017]). To investigate lineage and taxonomic differentiation further, variation was assessed using morphological data, including analysis of shape variation *via* geometric morphometrics (GM). Such analyses have proven to be useful tools for investigating morphological differentiation within species complexes (O'Leary *et al.*, 2016) and for providing supporting evidence in delimitation of new species (Martinez *et al.*, 2015).

2 | MATERIALS AND METHODS

Fishes were handled and euthanised prior to preservation in accordance with recommended guidelines for the use of fishes in research (Use of Fishes in Research Committee, 2014) and stress was ameliorated by minimising handling and through the use of the anaesthetic MS-222 for euthanasia.

2.1 | Specimen sampling

Following the morphological criteria outlined by Paugy (1986), specimens identified as *B. longipinnis* were sampled across a large portion of the species' distributional range, representing populations from drainage systems in Guinea, Liberia, Cameroon, Gabon, and the Rep. Congo (Figure 1 and Table 1). Tissue vouchers from additional populations from between Sierra Leone and Cameroon were unavailable for the current study. Eighty-two individuals were used to generate DNA barcodes (*col* sequences, 657 bp long; Table 1). In addition, 84 individuals from the same geographical localities were used to generate comparative morphological data (Supporting Information, Table S1). While most of the specimens sequenced for *col* were included in the morphometric assessment, there was not perfect overlap between the two datasets. This was due to the fact that some specimens used in the molecular analyses exhibited deformation during preservation and were not suitable for GM and similarly, tissue samples were not available for all formalin-fixed, alcohol-preserved specimens.

2.2 | DNA extraction, amplification and sequencing of *col* barcodes

Total genomic DNA was extracted from 95% ethanol-preserved tissue samples (*e.g.*, muscle, fin clips) using DNeasy Tissue Extraction Kit (Qiagen; www.qiagen.com) following the manufacturer's protocol. Amplification and sequencing of *col* barcodes were carried out using Folmer *et al.*'s (1994) universal primers LCO1490 (5'-GGTCAACAAATCATAA AGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3'). DNA amplification by PCR was performed in a 25 μl volume containing one Ready-To-Go PCR bead (GE Healthcare; www.gehealthcare.com), 21 μl of PCR-grade water, 1 μl of each primer (10 μM) and 2 μl of genomic DNA, under the following thermal profile: 5 min initial denaturation at 95°C, followed by 35 cycles of denaturation at 95°C for 60 s, annealing at 42°C for 60 s and extension at 72°C for 90 s, followed by a 7 min final extension at 72°C. Double-stranded PCR products were purified using AMPure (Agencourt, Beckman Coulter; www.beckman.com). Sequencing of each strand of amplified product was performed in a 5 μl volume containing 1 μl of primer (3.2 μM), 0.75 μl of BigDye Ready Reaction Mix (Gendx; www.gendx.com), 1 μl of BigDye buffer and 2.25 μl of PCR-grade water. Sequencing reactions consisted of a 2 min initial denaturation at 95°C, followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 45°C for 60 s and extension at

TABLE 1 Voucher specimens (catalogue, tissue numbers, locality data), *col* haplotype group, phenetic/phylogenetic group, and GenBank accession numbers for the *col* sequences generated in this study and included in the analyses

Catalog number	Tissue–voucher number	Country	Basin	Coordinates	*col* haplotype	Group (cluster/clade)	*col* GenBank accession number
AMNH 258028	LIB-003	Liberia	Sinoe	5° 4′ 6.4″ N, 8 ° 35′ 17.9″ W	1	I	MK592964
AMNH 263173	LIB-027	Liberia	Sinoe	5° 4′ 18.4″ N, 8 ° 33′ 12.2″ W	3	I	MK592944
AMNH 263206	LIB-113	Liberia	Sinoe	n/a	2	I	MK593006
AMNH 263206	LIB-114	Liberia	Sinoe	n/a	4	I	MK593007
AMNH 263206	LIB-115	Liberia	Sinoe	n/a	2	I	MK593008
AMNH 263215	LIB-138	Liberia	Sinoe	5° 6′ 30.3″ N, 8 ° 29 10.9″ W	6	I	MK593009
AMNH 263215	LIB-140	Liberia	Sinoe	5° 6′ 30.3″ N, 8° 29′ 10.9″ W	5	I	MK593010
AMNH 263215	LIB-144	Liberia	Sinoe	5° 6′ 30.3″ N, 8° 29′ 10.9″ W	2	I	MK593011
AMNH 263263	LIB-263	Liberia	Sinoe	5° 7′ 17.8″ N, 8° 30′ 3.5″ W	2	I	MK593012
AMNH 254202	9099 (AMCC 226378)	Guinea	Upper St. Pauls	7° 52′ 47.2″ N, 9° 04′ 58.2″ W	7	II	MK592956
AMNH 259205	AMCC 218453	Guinea	Upper St. Pauls	7° 46′ 57.8″ N, 9° 11′ 32.4″ W	7	II	MK592995
AMNH 257804	297 (AMCC 221837)	Guinea	Kolente	9° 51′ 43.1″ N, 12° 30′ 49″ W	9	III	MK592961
AMNH 257804	298 (AMCC 221838)	Guinea	Kolente	9° 51′ 43.1″ N, 12° 30′ 49″ W	10	III	MK592962
AMNH 257804	299 (AMCC 221839)	Guinea	Kolente	9° 51′ 43.1″ N, 12° 30′ 49″ W	10	III	MK592963
AMNH 260951	AMCC 220509	Guinea	Kolente	9° 57′ 47.1″ N, 12° 49′ 22″ W	9	III	MK592997
AMNH 263184	LIB-034	Liberia	Sinoe	5° 4′ 8.6″ N, 8° 29′ 56″ W	8	III	MK593001
AMNH 263184	LIB-037	Liberia	Sinoe	5° 4′ 8.6″ N, 8° 29′ 56″ W	11	III	MK593002
AMNH 263190	LIB-063	Liberia	Sinoe	5° 2′ 46.4″ N, 8° 33′ 15.1″ W	11	III	MK593003
AMNH 263201	LIB-093	Liberia	Sinoe	5° 5′ 31.0″ N, 8° 32′ 34.1″ W	12	III	MK593004
AMNH 263201	LIB-094	Liberia	Sinoe	5° 5′ 31.0″ N, 8° 32′ 34.1″ W	13	III	MK593005
AMNH 257800	AMCC 221600	Guinea	Kolente	9° 57′ 39.4″ N, 12° 48′ 38″ W	15	IV	MK592957
AMNH 257800	AMCC 221601	Guinea	Kolente	9° 57′ 39.4″ N, 12° 48′ 38″ W	15	IV	MK592958
AMNH 257802	AMCC 221625	Guinea	Kolente	9° 57′ 39.4″ N, 12° 49.′14″ W	15	IV	MK592959
AMNH 257802	AMCC 221626	Guinea	Kolente	9° 57′ 39.4″ N, 12° 49′0.14″ W	15	IV	MK592960
AMNH 260951	AMCC 220508	Guinea	Kolente	9° 57′ 47.1″ N, 12° 49′ 22″ W	15	IV	MK592996
AMNH 261781	AMCC 221612	Guinea	Kolente	9° 57′ 52.2″ N, 12° 49′ 36″ W	15	IV	MK592998
AMNH 254050	8970 (AMCC 230452)	Liberia	St. Pauls	7° 14′ 55.3″ N, 9° 16′ 28.1″ W	14	IV	MK592954
AMNH 254050	8971 (AMCC 230453)	Liberia	St. Pauls	7° 14′ 55.3″ N, 9° 16′ 28.1″ W	14	IV	MK592955
AMNH 258328	AMCC 211334	Rep. Congo	Coastal Plain	4° 34′ 0.1″ S, 11° 48′ 16.3″ E	25	V	MK592971
AMNH 258328	AMCC 211335	Rep. Congo	Coastal Plain	4° 34′ 0.1″ S, 11° 48′ 16.3″ E	20	V	MK592972
AMNH 258336	AMCC 211339	Rep. Congo	Coastal Plain	4° 37′ 58.2″ S, 11° 49′ 25.6″ E	22	V	MK592973
AMNH 258336	AMCC 211340	Rep. Congo	Coastal Plain	4° 37′ 58.2″ S, 11° 49′ 25.6″ E	23	V	MK592974
AMNH 258350	AMCC 211317	Rep. Congo	Coastal Plain	4° 20′ 57.3″ S, 11° 38′ 11.7″ E	27	V	MK592976
AMNH 258350	AMCC 211318	Rep. Congo	Coastal Plain	4° 20′ 57.3″ S, 11° 38′ 11.7″ E	27	V	MK592977
AMNH 258360 °	AMCC 211331	Rep. Congo	Coastal Plain	4° 19′ 13.4″ S, 11° 37′ 55″ E	25	V	MK592978
AMNH 258360	AMCC 211332	Rep. Congo	Coastal Plain	4° 19′ 13.4″ S, 11° 37′ 55″ E	21	V	MK592979
AMNH 258364	AMCC 211329	Rep. Congo	Coastal Plain	4° 19′ 32.8″ S, 11° 35′ 25.7″ E	20	V	MK592980
AMNH 258364	AMCC 211330	Rep. Congo	Coastal Plain	4° 19′ 32.8″ S, 11° 35′ 25.7″ E	27	V	MK592981
AMNH 258424	AMCC 211569	Rep. Congo	Coastal Plain	4° 21′ 29.1″ S, 11° 36′ 36.3″ E	26	V	MK592992
AMNH 258424	AMCC 211570	Rep. Congo	Coastal Plain	4° 21′ 29.1″ S, 11° 36′ 36.3″ E	26	V	MK592993
AMNH 258140	AMCC 211337	Rep. Congo	Kouilou	4° 8′ 44.4″ S, 11° 42′ 48.1″ E	25	V	MK592968
AMNH 258140	AMCC 211338	Rep. Congo	Kouilou	4° 8′ 44.4″ S, 11° 42′ 48.1″ E	25	V	MK592969
AMNH 258147	AMCC 211343	Rep. Congo	Kouilou	4° 1′ 59.6″ S, 11° 41′ 36.9″ E	25	V	MK592970
AMNH 258383	AMCC 211566	Rep. Congo	Kouilou	4° 2′ 1.1″ S, 11° 41′ 36.2″ E	25	V	MK592982

TABLE 1 (Continued)

Catalog number	Tissue–voucher number	Country	Basin	Coordinates	*col* haplotype	Group (cluster/clade)	*col* GenBank accession number
AMNH 258387	AMCC 211582	Rep. Congo	Kouilou	4° 2′ 1.1″ S, 11° 41′ 36.2″ E	25	V	MK592983
AMNH 258398	AMCC 211575	Rep. Congo	Kouilou	4° 6′ 54.4″ S, 11° 40′ 54.7″ E	25	V	MK592984
AMNH 258398	AMCC 211576	Rep. Congo	Kouilou	4° 6′ 54.4″ S, 11° 40′ 54.7″ E	25	V	MK592985
AMNH 258401	AMCC 211587	Rep. Congo	Kouilou	4° 5′ 44.7″ S, 11° 39′ 59.1″ E	25	V	MK592986
AMNH 258401	AMCC 211588	Rep. Congo	Kouilou	4° 5′ 44.7″ S, 11° 39′ 59.1″ E	25	V	MK592987
AMNH 258406	AMCC 211585	Rep. Congo	Kouilou	4° 7′ 45.5″ S, 11° 43′ 2.1″ E	19	V	MK592988
AMNH 258406	AMCC 211586	Rep. Congo	Kouilou	4° 7′ 45.5″ S, 11° 43′ 2.1″ E	19	V	MK592989
AMNH 258452	AMCC 211589	Rep. Congo	Kouilou	4° 14′ 27.2″ S, 11° 41′ 35.7″ E	19	V	MK592994
AMNH 258125	AMCC 211326	Rep. Congo	Lake Koubambi (Kouilou)	4° 9′ 0.6″ S, 11° 44′ 41.8″ E	19	V	MK592965
AMNH 258125	AMCC 211327	Rep. Congo	Lake Koubambi (Kouilou)	4° 9′ 0.6″ S, 11° 44′ 41.8″ E	19	V	MK592966
AMNH 258134	AMCC 211336	Rep. Congo	Lake Yangala (Noumbi)	4° 7′ 26.1″ S, 11° 35′ 48.9″ E	25	V	MK592967
AMNH 258139	AMCC 211320	Rep. Congo	Lake Yangala (Noumbi)	4° 7′ 25.8″ S, 11° 35′ 48.7″ E	25	V	MK592942
AMNH 258139	AMCC 211321	Rep. Congo	Lake Yangala (Noumbi)	4° 7′ 25.8″ S, 11° 35′ 48.7″ E	25	V	MK592943
AMNH 258413	AMCC 211578	Rep. Congo	Lake Youbi (Kouilou)	4° 11′ 10.1″ S, 11° 39′ 58″ E	24	V	MK592990
AMNH 258413	AMCC 211579	Rep. Congo	Lake Youbi (Kouilou)	4° 11′ 10.1″ S, 11° 39′ 58″ E	25	V	MK592991
AMNH 253900	8720	Rep. Congo	Niari	3° 51′ 58.4″ S, 12° 20′ 47.2″ E	28	V	MK592951
AMNH 253900	8721	Rep. Congo	Niari	3° 51′ 58.4″ S, 12° 20′ 47.2″ E	28	V	MK592952
OS 19397	BLS14-521	Gabon	Ogowe	0° 46′ 21.11″ S, 12° 54′ 42.19″ E	29	V	MK593021
OS 19871	BLS14-122	Gabon	Ogowe	0° 59′ 37.88″ S, 13° 31′ 34.64″ E	29	V	MK593023
AMNH 249831	6310 (AMCC 257287)	Cameroon	Sanaga	3° 49′ 12.25″ N, 10° 8′ 9.72″ E	18	V	MK592948
AMNH 249551	6257 (AMCC 257637)	Cameroon	Wouri	3° 59′ 52.33″ N, 9° 50′ 57.97″ E	17	V	MK592945
AMNH 249561	6258 (AMCC 257650)	Cameroon	Wouri	4° 27′ 17.29″ N, 9° 58′ 44.48″ E	16	V	MK592946
AMNH 249561	6259 (AMCC 257651)	Cameroon	Wouri	4° 27′ 17.29″ N, 9° 58′ 44.48″ E	18	V	MK592947
AMNH 263041	AMCC 227162	Rep. Congo	Loeme	4° 35′ 3.2″ S, 12° 6′ 3.5″ E	32	VI	MK592999
AMNH 263130	AMCC 227302	Rep. Congo	Loeme	4° 36′ 32″ S, 12° 8′ 57.1″ E	33	VI	MK593000
AMNH 263477	AMCC 230238	Rep. Congo	Loeme	4° 39′ 32.9″ S, 12° 10′ 21″0.2 E	30	VI	MK593013
AMNH 263489	AMCC 230252	Rep. Congo	Loeme	4° 33′ 39.5″ S, 12° 6 ′30.4″ E	31	VI	MK593014
AMNH 263489	AMCC 230253	Rep. Congo	Loeme	4° 33′ 39.5″ S, 12° 6′ 30.4″ E	34	VI	MK593015
OS 19394	BLS14-040	Gabon	Ogowe	0° 50′ 32.82″ S, 12° 57′ 44.96″ E	35	VII	MK593018
OS 19395	BLS14-041	Gabon	Ogowe	0° 50′ 32.82″ S, 12° 57′ 44.96″ E	35	VII	MK593019
AMNH 253875	8865 (AMCC 236034)	Rep. Congo	Kouilou	4° 4 48′0.5″ S, 12° 8′ 18.4″ E	36	VIII	MK592949
AMNH 253875	8866 (AMCC 236035)	Rep. Congo	Kouilou	4° 4 48′0.5″ S, 12° 8′ 18.4″ E	36	VIII	MK592950
AMNH 258346	AMCC 211333	Rep. Congo	Kouilou	4° 26 33′0.1″ S, 11° 45′ 43.6″ E	36	VIII	MK592975
AMNH 253934	8784	Rep. Congo	Niari	4° 6′ 01″ S, 13° 03′ 38″ E	36	VIII	MK592953
OS 19392	BLS14-047	Gabon	Ogowe	0° 50′ 32.82″ S, 12° 57′ 44.96″ E	37	VIII	MK593016
OS 19393	BLS14-037	Gabon	Ogowe	0° 50′ 32.82″ S, 12° 57′ 44.96″ E	37	VIII	MK593017
OS 19396	BLS14-031	Gabon	Ogowe	0° 50′ 32.82″ S, 12° 57′ 44.96″ E	37	VIII	MK593020
OS 19398	BLS14-042	Gabon	Ogowe	0° 50′ 32.82″ S, 12° 57′ 44.96″ E	37	VIII	MK593022

AMCC: Ambrose Monell CryoCollection (AMNH); AMNH: American Museum of Natural History; OS: Oregon State University.

72°C for 4 min, followed by a 3 min final extension at 72°C. All sequencing reactions were purified using CleanSEQ (Agencourt) and electrophoresed on an Applied Biosystems 3700 automated DNA sequencer (www.appliedbiosystems) in the American Museum of Natural History (AMNH) molecular systematics laboratories. Contig assembly and sequence editing were performed using the software Geneious 11.0.2 (Kearse *et al*., 2012).

2.3 | Assessment of *col* sequence variation and estimation of genealogical relationships

In an effort to rule out potential specimen misidentification, tissue mislabelling, or tissue/extract contamination issues, all barcodes generated were compared with reference sequences deposited in GenBank using NCBI basic local alignment selection tool (BLAST; Altschul *et al*., 1990; Johnson *et al*., 2008) and the best match was taken as the closest estimate of taxonomic identity. Prior to analyses of DNA sequence variation, *col* sequences were aligned using the software MUSCLE (Edgar, 2004) under default parameters. Two matrices of pairwise genetic distances were computed from the aligned sequence data using the R (www.r-project.org) package *ape* (Paradis *et al*., 2004; Popescu *et al*., 2012): one indicating absolute (uncorrected) distances (*p*-distances) and another indicating corrected (model-based) genetic distances. The latter was computed based on the TN93 substitution model, which best fit the *col* data (of 24 models evaluated) according to the Bayesian information criterion (BIC) as implemented in jModelTest (Posada, 2008) under the following likelihood settings: number of substitution schemes = 3; base frequencies = +*F*; rate variation = +*I* and + *G* with nCat = 4; and base tree for likelihood calculations = fixed BIONJ-JC. The resulting corrected distance matrix was used as input to generate an unrooted tree (network) and a rooted tree (dendrogram) based on the neighbour joining (NJ; Saitou & Nei, 1987) and the unweighted pair group method with arithmetic mean (UPGMA; Sokal & Michener, 1958) hierarchical clustering algorithms, respectively. Both NJ and UPGMA analyses were performed with the *phangorn* package in R (Schliep, 2010). The number and identity of haplotypes was estimated from the aligned sequence data using the *pegas* package in R (Paradis, 2010). A haplotype network displaying relationships among *col* haplotypes was inferred using the minimum joining network method (Bandelt *et al*., 1999) as implemented in the software PopArt (Leigh & Bryant, 2015), based on an absolute distance matrix calculated for all pairwise comparisons of haplotypes. In addition to distance-based methods for assessment of genetic variation and relationships among alleles, a character-based, phylogenetic approach, maximum likelihood (ML) in this case, was used to infer evolutionary relationships among samples. Phylogeny estimation *via* ML was carried out on the aligned *col* data using RAxMl 7.2.8 Black Box (Stamatakis, 2006). Based on the findings of Arroyave and Stiassny (2011), the alestid species *Micralestes acutidens* (Peters 1852) (AMNH 239476; GenBank Accession Number JF800942) was used as outgroup for the phylogenetic analysis.

2.4 | Assessment of phenotypic variation

GM was used to compare body shape variation in *B. longipinnis* from across its geographic range. The aim was to identify potentially diagnostic features for samples from phylogeographic lineages revealed by *col* data that principally corresponded to two major freshwater ecoregions/ichthyofaunal provinces of continental Africa: upper Guinea (UG) and lower Guinea (LG) (Abell *et al*., 2008; Roberts, 1975). Eighty-four individuals were photographed in a standardised lateral view, including 50 fish from UG and 34 from LG (Supporting Information, Table S1). Sixteen homologous landmarks were digitised in tpsDIG2 (Rohlf, 2015) and were used to describe body shape variation in *B. longipinnis*. Some specimens displayed upward or downward arching of the body as artefacts of preservation. The unbending function in tpsUtil (Rohlf, 2015) was used to remove this source of variation, fitting a quadratic curve to reference landmarks placed along a true linear baseline and adjusting or unbending other landmarks, accordingly (Valentin *et al*., 2008). For this, two temporary landmarks along the lateral midline of the body were added, one at the base of the caudal fin and another directly above the pelvic fin. These landmarks were removed after unbending. Shapes were aligned in the R package *geomorph* (Adams & Otarola-Castillo, 2013), using generalised Procrustes analysis. Overall patterns of shape variation were visualised with a principle component analysis paired with thin-plate spline analysis to generate warp grids for displaying shape deformations at extremes of principle component (PC) axes relative to mean shape.

Before making comparisons between geographic locations, the effects of size on body shape (*i.e*., allometry), which could potentially confound results, needed to be removed. Due to the presence of sexual dimorphism in the examined specimens, a 10,000 permutation multivariate analysis of covariance (MANCOVA) with shape grouped by sex and log-centroid size as the covariate was first used to test whether males and females had similar allometric slopes, such that a single function could be used to adjust for size. Next, regressions of shape on log-centroid size were performed, taking the residuals as size-corrected shape data (Martinez & Sparks, 2017; McCord & Westneat, 2016). A MANOVA, with significance based on 10,000 permutations, was used to compare whether morphologies of allometrically-adjusted shape data differed between UG and LG. Finally, region-specific shape variation of specimens, as indicated by warp grids of group means, was used to identify linear dimensions for further comparisons, using ANOVAs.

3 | RESULTS

3.1 | Geographic patterns of *col* divergence in *B. longipinnis*

For all *col* sequences, the best nucleotide–nucleotide basic local-alignment search tool (BLASTn) match corresponded to *Bryconalestes*, probably ruling out major mislabelling, or contamination issues. With respect to species identity, only 54 of the 82 *col* sequences had *B. longipinnis* as top BLAST match; the remaining 28 samples closer aligned with *B. bartoni* (21) and *B. tholloni* (7), respectively. From the

sample of 82 individuals genotyped, a total of 37 *coI* haplotypes were identified (Table 1), 19 of which are singletons. Absolute pairwise genetic distances among *coI* haplotypes range from 0.15% to 15.22%, indicating high genetic differentiation within and between haplotypes from different, or the same, basins and ichthyofaunal provinces. Absolute pairwise genetic distances among samples–sequences from the same country are also high: Cameroon (≤4.11%), Gabon (≤11.06%), Guinea (≤6.85%), Liberia (≤8.68%) and Rep. Congo (≤10.5%). No premature stop codons, frameshifts, insertions, deletions, or heterozygotes were detected, suggesting that the high levels of *coI* divergence are unlikely to be artefacts of having inadvertently sequenced nuclear mitochondrial DNA (mtDNA) segments (NUMT; Song *et al*., 2008) or nuclear pseudogenes. Patterns of overall similarity among haplotypes (NJ tree) and among samples/sequences (UPGMA tree) are presented in Figure 3. A haplotype network portraying genealogical relationships among haplotypes is presented in Figure 4. Phylogenetic relationships (ML gene tree) among sampled individuals as inferred from *coI* sequence data are presented in Figure 5. Both distance-based and phylogenetic analytical methods identify eight clusters/clades (I–VIII) that diverge by more than 4% (*p*-distance) and therefore potentially represent cryptic species (Figures 3 and 5). All methods identify two major genetically distinct groups, *c*. 18% divergent (TN93 corrected distance), corresponding with two identifiable ichthyofaunal provinces: upper and lower Guinea (Figures 3–5).

3.2 | Geographic patterns of phenotypic variation in *B. longipinnis*

A principle component analysis of body shape in the sampled specimens of *B. longipinnis* reveals that the largest axis of diversity (PC 1) is characterised mainly by differences in body depth, which are related to sexual dimorphism (Figure 6). Females generally possess more slender, streamlined bodies, whereas males tend to be deeper-bodied. Along PC 2, fish from the different ichthyofaunal provinces (UG and LG) overlap broadly at low values, but individuals from LG are found at the largest scores on this axis and those from UG occupy the lowest scores. However, there is no evidence of clear morphological separation based on geography for subsequent PCs. Across all specimens evaluated, the test of slopes from a MANCOVA shows a significant interaction between sex and size ($P < 0.01$), indicating that separate allometric adjustments should be made on males and females (see Supporting Information, Table S2 for regression results). In both males and females, significant but weak differences in body shape between UG and LG ecoregions were found. In males, regional differences account for 12.5% of overall body shape variance ($F_{1,34} = 4.84$, $P < 0.001$) and only 4.6% in females ($F_{1,46} = 2.24$, $P < 0.05$). Within each sex, visualisation of shape changes with thin-plate spline analyses (Supporting Information, Figure S1) reveals a number of morphological features that potentially differ between the upper and lower Guinean regions. In females, the depth of the caudal peduncle relative to its length is significantly higher in fishes from UG (Figure 7). The same was true for female eye diameter. In males, differences in morphological features were apparent, with relative peduncle depth and anal and dorsal-fin base lengths being greater in fish from UG. In contrast, snout length was significantly higher in males from LG than from UG.

4 | DISCUSSION

This study represents the first assessment of genetic and morphometric variation in the alestid *B. longipinnis* across its extent of occurrence. *coI* variation in samples from across the putative species' geographic range is considerably greater than the traditionally employed *c*. 3% sequence divergence heuristic threshold for conspecifics (Avise, 2000; Herbert *et al*., 2003, 2004). Despite all sampled individuals conforming to the morphological criteria outlined by Paugy (1986) for identification of *B. longipinnis*, only *c*. 66% of the samples top BLAST matched with *B. longipinnis*, while the remaining *c*. 34% with closely related species *B. bartoni* and *B. tholloni*. This finding, however, is not entirely surprising considering that taxonomic identifications of vouchers in online biodiversity repositories such as GenBank are not necessarily completely reliable, especially for highly diverse and taxonomically difficult groups (Nilsson *et al*., 2006; Vilgalys, 2003).

Distance-based (Figure 3), network (Figure 4) and phylogenetic (Figure 5) analyses of *coI* variation all converge on significant genetic differentiation across the range of the species, distinguishing eight clusters/clades (I–VIII) that diverge by more than 4% (*p*-distance) and are herein regarded as operational taxonomic units (OTU) representing potential cryptic species of a *B. longipinnis* species complex. Furthermore, all methods reveal a clear pattern of divergence, in which two main clusters/clades corresponding to distinct ichthyofaunal provinces are identified: an upper Guinean (UG; Liberia, Guinea) and a lower Guinean (LG; Cameroon, Gabon, Rep. Congo) grouping, a result that supports the earlier findings of Paugy (1982). However, within these groupings there seems to be no strong association of haplotypes with countries or river drainages (Table 1). Within the UG group for example, samples/haplotypes from river basins across Guinea are interspersed with those from Liberia. Similarly, within the LG group, samples/haplotypes from Gabon are scattered throughout basins from the Rep. Congo, although samples from Cameroonian drainages do appear to be distinct and more similar/closely related to each other than to the remaining LG samples/haplotypes. The fact that sympatric samples do not always cluster together but instead appear to be more closely related to samples from other basins suggests that, at smaller geographic scales such as basin and sub-basin, genetic differentiation in *B. longipinnis* is not explained by geography. Despite this lack of phylogeographic signal at local scales, there is a remarkably high degree of *coI* sequence divergence within most geographic units, again, greater than the empirical threshold values standardly expected for conspecifics.

The syntypical series of *Brachyalestes longipinnis* (BMNH 1864.3.16:2-7) is recorded as being from 'Sierra Leone' (Paugy, 1986) and while it was possible to examine it in the present study, these specimens were in relatively poor condition and unsuitable for morphometric analyses. Similarly, it was impossible to obtain DNA

FIGURE 3 Legend on next page.

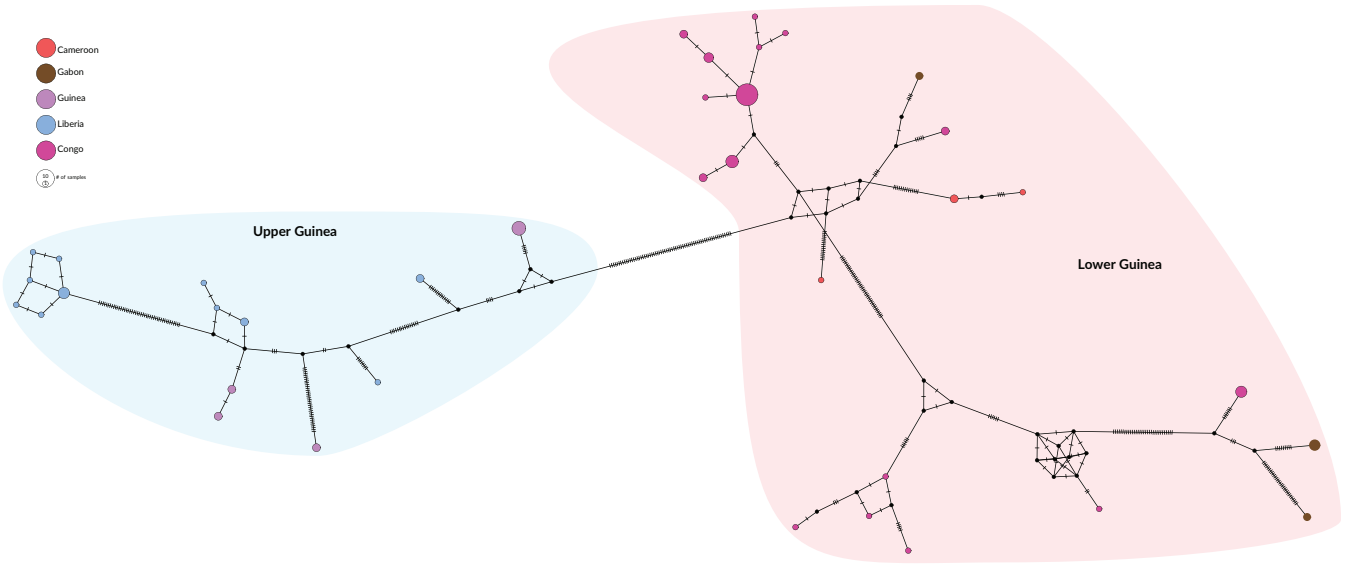

FIGURE 4 Haplotype network portraying genealogical relationships among *coI* haplotypes of *Bryconalestes longipinnis*

sequence data from them for inclusion in genetic analyses. Consequently, it was not feasible to assign the types to any of the haplotype lineages identified in the current study.

Assuming that *B. longipinnis* indeed corresponds to a monophyletic assemblage independent from other *Bryconalestes* spp. and that all specimens sampled in this study are members of that assemblage, then the discovery of extremely high levels of *coI* sequence variation between putative conspecifics strongly suggests the presence of cryptic diversity. These findings therefore imply that current taxonomy represents a significant underestimation of species-level diversity within the group. The discovery of high levels of *coI* divergence in *B. longipinnis* suggestive of cryptic diversity is not entirely surprising, considering that this is a widespread, taxonomically understudied and poorly diagnosed species. Certainly, a common shortcoming of traditional taxonomic studies based solely on phenetic analysis of morphological variation, is to mistakenly regard as conspecific morphologically similar but genetically (and ultimately evolutionarily) distinct populations (Dayrat, 2005). Such failure to recognise the existence of divergent lineages necessarily results in an underestimation of diversity; a situation to which widespread, morphologically homogeneous species of taxonomically neglected groups are particularly prone (Gill & Kemp, 2002; Goodier *et al*., 2011; Pyron & Burbrink, 2009; Manthey *et al*., 2011).

Other widely distributed African freshwater fishes with similar taxonomic uncertainty have been shown to represent species complexes when examined more closely and with broad geographic sampling (Snoeks *et al*., 2011). Similarly, recent DNA barcoding surveys aimed at assessing traditional morphology-based species identifications of fishes from the Congo basin (Decru *et al*., 2016) documented high levels of genetic divergence in several species, including the alestids *Hydrocynus vittatus* Castelnau 1861, *Brachypetersius altus* (Boulenger 1899) and *Brycinus imberi* (Peters 1852). Morphometric examination of voucher specimens of *H. vittatus*, however, did not reveal differences between genetically divergent groups, implying the presence of cryptic diversity (Decru *et al*., 2016). As in *H. vittatus* and *P. buchholzi* (Lavoué *et al*., 2010), there appears to be a mismatch between levels of observed genetic and morphological divergence within *B. longipinnis*. Continuing this trend, a more recent study that barcoded almost 200 species of African freshwater fishes from the Congo Basin and lower Guinea regions (Sonet *et al*., 2018) reported high levels of intraspecific *coI* divergence and therefore inconsistencies between morphology-based identifications and DNA barcode clustering. Most of these inconsistencies were attributed to undescribed taxonomic diversity, just as in the case of *B. longipinnis* reported here, and resulted in the proposal of 17 putative new species (Sonet *et al*., 2018). Likewise, similar to the findings of the present study, Sonet *et al*. (2018) reported that, in most cases, samples corresponding to the same morphospecies but from adjacent ichthyofaunal provinces showed considerable levels of *coI* divergence (up to 9.4%) while remaining morphologically uniform. This pattern highlights the importance of DNA barcoding studies and surveys aimed at preliminarily testing species limits in widespread species, especially when distributed across different biogeographic regions.

Body shape analyses based on geometric morphometrics were focused on the identification of morphological differences in *B. longipinnis* across its geographic range. Specifically, between populations from upper and lower Guinean biogeographic regions,

FIGURE 3 (a) Patterns of overall similarity among haplotypes (neighbour-joining tree) and (b) among samples/sequences (unweighted pair group method with arithmetic mean (UPGMA) tree) of *Bryconalestes longipinnis*. Horizontal axis in the UPGMA tree (dendrogram) indicates *coI* genetic divergence between samples and clusters (%) corrected (TN93) genetic distance

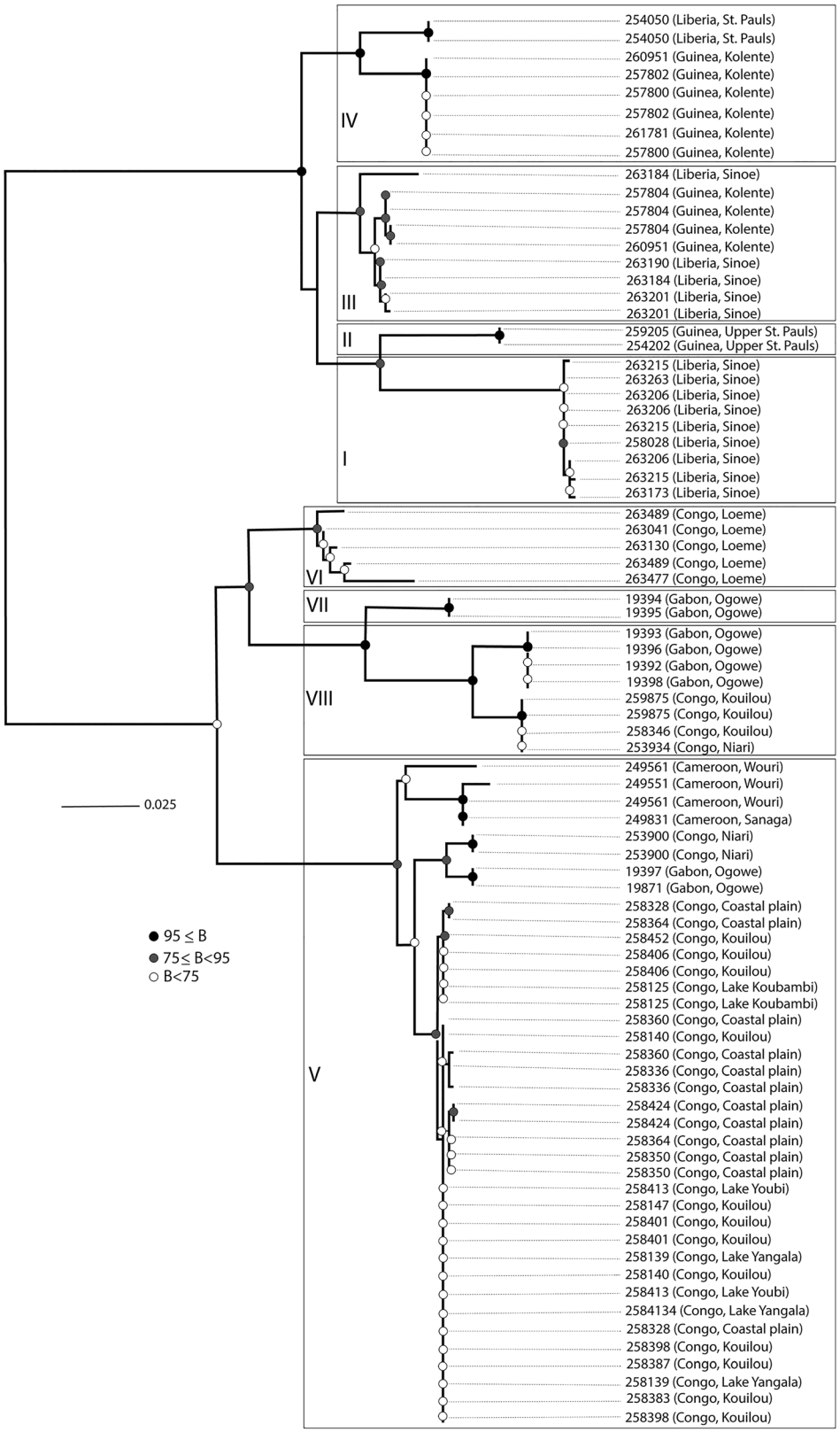

FIGURE 5 Phylogenetic relationships (maximum likelihood gene tree) among sampled individuals of *Bryconalestes longipinnis* as inferred from *col* sequence data. I–VIII, The eight clusters–clades diverge by more than 4% (*p*-distance) and therefore potentially represent cryptic species

which represented the primary divergence of lineages recovered from *col* sequence data. Consistent with previous work using meristics (Paugy, 1982, 1986), the present study found that shape data and linear measurements displayed considerable overlap between UG and LG populations and while statistically significant differences were observed, the ability to reliably utilise external body shape for regional delimitation remains elusive. It is interesting, however, that shape variation due to sexual dimorphism was considerably more prominent

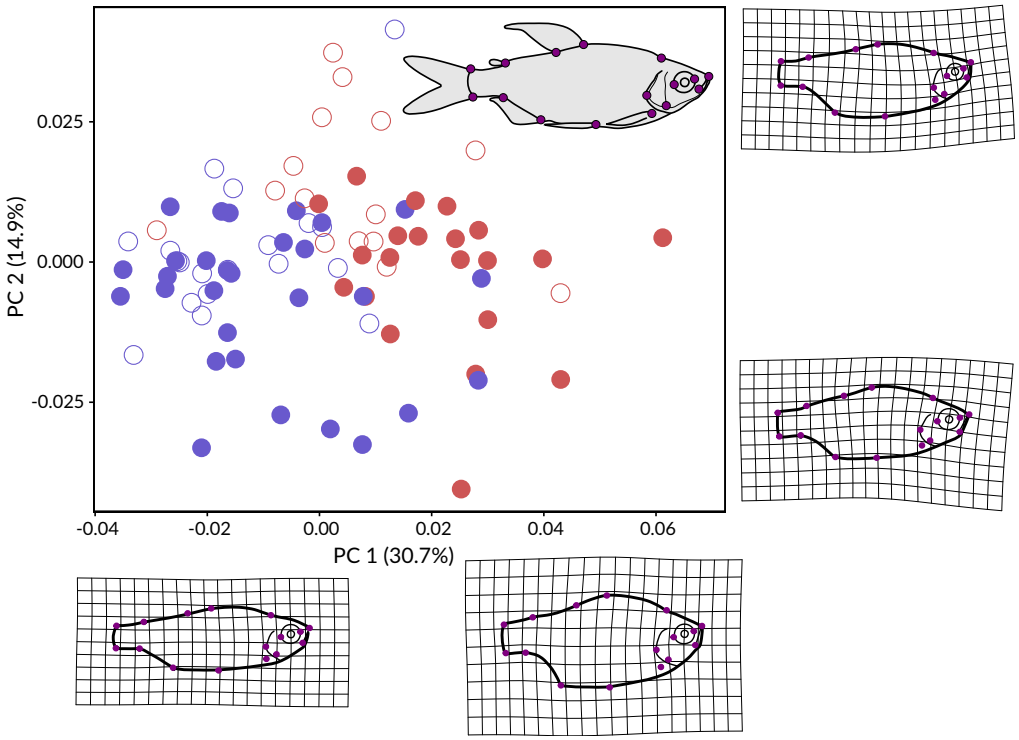

FIGURE 6 Principle components (PC) 1 and 2 of body shape variation in *Bryconalestes longipinnis*. (●), Males; (●), females. Specimens from upper and lower Guinea represented by closed circles and open circles, respectively. Landmarks used in this study (—●—) are shown in upper right of plot and warp grids display shape change at PC extremes, relative to mean body shape

than that between geographic lineages of *B. longipinnis*, but such a pattern is not without precedent. It is thought that under some conditions, the processes leading to strong dimorphism could be antagonistic to those promoting disruptive selection and eventually, speciation (Bolnick & Doebeli, 2003). Regardless, a more extensive and detailed evaluation of other morphological traits beyond external body shape (*e.g.*, meristic characters, skeletal anatomy) appears necessary to establish whether compelling diagnostic morphological features do exist between the divergent lineages of the *B. longipinnis* complex uncovered in the present study.

The limitations of using a single mtDNA marker in taxonomic and phylogeographic studies have been discussed extensively, and introgressive hybridisation (among other evolutionary processes) represents a possible source of error. mtDNA introgression resulting from the hybridisation of *B. longipinnis* with other *Bryconalestes* spp. could explain some of the observed elevated levels of intraspecific *coI* divergence in the absence of profound morphological differentiation. On the other hand, if the high level of cryptic diversity suggested by the current study is real, hybridisation with differential introgression between cryptic *B. longipinnis* lineages could be effectively blurring morphological boundaries while maintaining elevated levels of *coI* divergence between them. At this stage, however, such hybridisation scenarios are speculative and further research is needed to properly test these hypotheses.

In conclusion, this study fails to corroborate the hypothesis that *B. longipinnis* is, as implied by currently employed morphological criteria, a single, widespread species. Instead the results support the existence of a species complex of phenetically similar but genetically and evolutionarily distinct entities, in which two major phylogeographic lineages, encompassing the grouping of upper and lower Guinean populations, can be distinguished. However, even within each of these main geographical regions, levels of *coI* sequence divergence suggest that there may be as many as eight cryptic species encompassed within the current concept of *B. longipinnis*. While it is important to emphasise that DNA barcoding is primarily a method designed to identify described species and to provisionally recognise undescribed ones, not to define them (Witt *et al.*, 2006), the current study provides a much-needed starting point for future revisional studies. The phylogeographic structure revealed herein provides a framework that facilitates the targeted sampling of exemplars for detailed morphological analyses necessary to properly diagnose species, define species limits, and ultimately clarify the taxonomy of the group. Thus, this study underscores the utility of DNA barcoding as a tool for rapidly assessing genetic intraspecific variation that could lead to the discovery of cryptic diversity, either in the form of unreported phylogeographic lineages or in the form of cryptic, undescribed species.

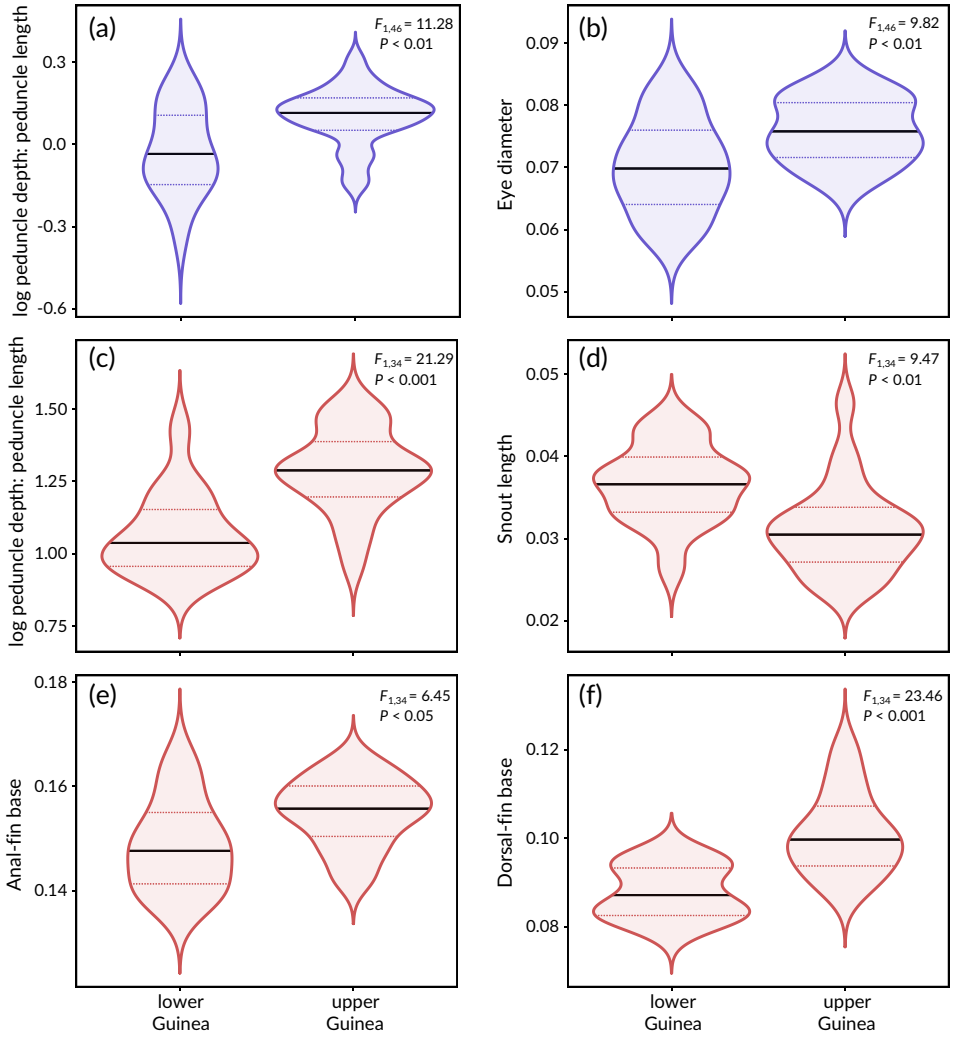

FIGURE 7 Violin plots displaying variation in linear measurements made on *Bryconalestes longipinnis* from upper and lower Guinea. (——), Median values; (·····), first and third quantiles. Comparisons in females include: (a) the ratio of caudal peduncle depth to peduncle length and (b) eye diameter. Male comparisons were made on the: (c) caudal-peduncle depth-to-length ratio, (d) snout length, (e) length of anal fin base, and (f) length of dorsal fin base. Results from ANOVAs between regions are included in upper right corners of panels. Note that linear measurements were calculated from aligned and size-adjusted landmark data, with units in relative scale (See Figure S1 for justification and diagram of linear measurements.)

ACKNOWLEDGEMENTS

Our thanks to Brian Sidlauskas (Oregon State University) for the loan of samples from Gabon and to Gina Walsh (University of Witswatersrand) and Victor Mamonekene (University of Marien Ngouabi) for their collecting efforts in the Rep. Congo.

Contributions

J.A. and M.L.J.S. conceived and designed the study. M.L.J.S. and J.A. collected voucher specimens in the field. M.L.J.S. conducted morphological taxonomic determination. J.A. generated and analysed comparative molecular data. C.M.M. carried out geometric morphometrics analyses. J.A. drafted the initial version of the manuscript. All authors read, edited, enhanced and approved the final version of the manuscript.

ORCID

Jairo Arroyave https://orcid.org/0000-0002-5389-1789
Melanie L. J. Stiassny https://orcid.org/0000-0001-8220-4768

REFERENCES

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., … Petry, P. (2008). Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *Bioscience*, *58*, 403–414.

Adams, D. C., & Otarola-Castillo, E. (2013). Geomorph: An R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, *4*(4), 393–399.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*, 403–410.

Arroyave, J., & Stiassny, M. L. J. (2011). Phylogenetic relationships and the temporal context for the diversification of African characins of the family Alestidae (Ostariophysi: Characiformes): Evidence from DNA sequence data. *Molecular Phylogenetics and Evolution*, *60*, 385–397.

Avise, J. C. (2000). *Phylogeography. The history and formation of species*. Cambridge, MA: Harvard University Press.

Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, *16*(1), 37–48.

Bolnick, D. I., & Doebeli, M. (2003). Sexual dimorphism and adaptive speciation: Two sides of the same ecological coin. *Evolution*, *57*(11), 2433–2449.

Calcagnotto, D., Schaefer, S. A., & DeSalle, R. (2005). Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. *Molecular Phylogenetics and Evolution*, *36*, 135–153.

Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, *87*, 407–415.

Decru, E., Vreven, E., & Snoeks, J. (2012). A revision of the West African *Hepsetus* (Characiformes: Hepsetidae) with a description of *Hepsetus akawo* sp. nov. and a redescription of *Hepsetus odoe* (Bloch, 1794). *Journal of Natural History*, *46*, 1–23.

Decru, E., Vreven, E., & Snoeks, J. (2013). A revision of the Lower Guinean *Hepsetus* species (Characiformes; Hepsetidae) with the description of *Hepsetus kingsleyae* sp. nov. *Journal of Fish Biology*, *82*(4), 1351–1375.

Decru, E., Snoeks, J., & Vreven, E. (2015). Taxonomic evaluation of the *Hepsetus* from the Congo Basin with the revalidation of *H. microlepis* (Teleostei: Hepsetidae). *Ichthyological Exploration of Freshwaters*, *26*, 273–287.

Decru, E., Moelants, T., De Gelas, K., Vreven, E., Verheyen, E., & Snoeks, J. (2016). Taxonomic challenges in freshwater fishes: A mismatch between morphology and DNA barcoding in fish of the north-eastern part of the Congo basin. *Molecular Ecology Resources*, *16*(1), 342–352.

Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*(1), 113.

Eschmeyer, W.N., Fricke, R., & van der Laan, R. (2018). Catalog of fishes: genera, species. Retrieved from https://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp

Folmer, O., Hoeh, W. R., Black, M. B., & Vrijenhoek, R. C. (1994). Conserved primers for PCR amplification of mitochondrial DNA from different invertebrate phyla. *Molecular Marine Biology and Biotechnology*, *3*, 294–299.

Géry, J., & Mahnert, V. (1977). Description d'une espèce et d'une sous-espèce nouvelles d'Afrique occidentale du genre *Brycinus* (Pisces, Cypriniformes, characoidei), avec une revue de la supra-espèce *Brycinus longipinnis* (Günther, 1864). *Revue Suisse de Zoologie*, *84*(3), 617–643.

Gill, A., & Kemp, J. M. (2002). Widespread Indo-Pacific shore fish species: A challenge for taxonomists, biogeographers, ecologists and fisheries and conservation managers. *Environmental Biology of Fishes*, *65*, 165–174.

Goodier, S. A. M., Cotterill, F. P. D., O'Ryan, C., Skelton, P. H., & de Wit, M. J. (2011). Cryptic diversity of African Tigerfish (genus *Hydrocynus*) reveals palaeogeographic signatures of linked neogene geotectonic events. *PLoS One*, *6*(12), e28775.

Günther, A. (1864). *Catalogue of the fishes in the British Museum. Catalogue of the physostomi, containing the families siluridae, characinidae, haplochitonidae, sternoptychidae, scopelidae, stomiatidae in the collection of the British Museum*. v. 5: i-xxii + 1–455. London: British Museum (Natural History).

Hebert, P. D., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B*, *270*(Suppl 1), S96–S99.

Hebert, P. D., Stoeckle, M. Y., Zemlak, T. S., & Francis, C. M. (2004). Identification of birds through DNA barcodes. *PLoS Biology*, *2*(10), e312.

Hoedeman, J. J. (1951). Studies on African characid fishes I. The tribe Alestidi (I). *Beaufortia*, *3*, 1–8.

Hubert, N., Bonillo, C., & Paugy, D. (2005). Early divergence among the Alestidae (Teleostei, Ostariophyses, Characiformes): Mitochondrial evidences and congruence with morphological data. *Comptes Rendus Biologies*, *328*, 477–491.

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, E., Burridge, M., … Bernatchez, L. (2008). Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One*, *3*, e2490.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: A better web interface. *Nucleic Acids Research*, *36*(suppl_2), W5–W9.

Joyce, D. A., Lunt, D. H., Genner, M. J., Turner, G. F., Bills, R., & Seehausen, O. (2011). Repeated colonization and hybridization in Lake Malawi cichlids. *Current Biology*, *21*, R108–R109.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Thierer, T. (2012). Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649.

Lalèyè, P., & Moelants, T. (2010). *Brycinus longipinnis*. The IUCN Red List of Threatened Species. Version 2014.3. www.iucnredlist.org

Lara, A., Ponce de León, J. L., Rodriguez, R., Casane, D., Cote, G., Bernatchez, L., & García-Machado, E. R. (2010). DNA barcoding of Cuban freshwater fishes: Evidence for cryptic species and taxonomic conflicts. *Molecular Ecology Resources*, *10*(3), 421–430.

Lavoué, S., Miya, M., Arnegard, M. E., McIntyre, P. B., Mamonekene, V., & Nishida, M. (2010). Remarkable morphological stasis in an extant vertebrate despite tens of millions of years of divergence. *Proceedings of the Royal Society B*, *278*(1708), 1003–1008.

Leigh, J. W., & Bryant, D. (2015). PopArt: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116.

Lowenstein, J. H., Osmundson, T. W., Becker, S., Hanner, R., & Stiassny, M. L. J. (2011). Incorporating DNA barcodes into a multi-year inventory of the fishes of the hyperdiverse lower Congo River, with a multi-gene performance assessment of the genus Labeo as a case study. *Mitochondrial DNA*, *22*, 52–70.

Manthey, J. D., Klicka, J., & Spellman, G. M. (2011). Cryptic diversity in a widespread North American songbird: Phylogeography of the Brown Creeper (*Certhia americana*). *Molecular Phylogenetics and Evolution*, *58*, 502–512.

Martinez, C. M., & Sparks, J. S. (2017). Malagasy cichlids differentially limit impacts of body shape evolution on oral jaw functional morphology. *Evolution*, *71*(9), 2219–2229.

Martinez, C. M., Arroyave, J., & Sparks, J. S. (2015). A new species of *Ptychochromis* from southeastern Madagascar (Teleostei: Cichlidae). *Zootaxa*, *4044*(1), 79–92.

McCord, C. L., & Westneat, M. W. (2016). Evolutionary patterns of shape and functional diversification in the skull and jaw musculature of triggerfishes (Teleostei: Balistidae). *Journal of Morphology*, *277*(6), 737–752.

Nilsson, R. H., Ryberg, M., Kristiansson, E., Abarenkov, K., Larsson, K.-H., & Kõljalg, U. (2006). Taxonomic reliability of DNA sequences in public sequence databases: A fungal perspective. *PLoS One*, *1*(1), e59.

Norris, S. M. (2002). A revision of the African electric catfishes, family Malapteruridae (Teleostei, Siluriformed), with erection of a new genus and description of fourteen new species and an annotated bibliography. *Annals Royal Museum for Central Africa, Zoological Sciences*, *289*, 1–155.

O'Leary, S. J., Martinez, C. M., Baumann, H., Abercrombie, D. L., Conover, D. O., Poulakis, G. R., … Chapman, D. D. (2016). Population genetics and geometric morphometrics of the key silverside, *Menidia conchorum*, a marine fish in a highly-fragmented, inland habitat. *Bulletin of Marine Sciences*, *92*(1), 33–50.

Paugy, D. (1982). Mise en synonymie d' *Alestes chaperi* (Sauvage 1882). Avec. *A. longipinnis* (Gunther 1864). (Pisces, Characidae). *Cybium*, *6*, 75–90.

Paugy, D. (1986). *Révision systématique des Alestes et Brycinus Africains (Pisces, Characidae)* (Collection Études et Théses). Éditions de L'ORSTOM, p. 295. Paris: IRD Editions.

Paugy, D., & Schaefer, S. A. (2007). Alestidae. In M. L. J. Stiassny, G. G. Teugels, & C. D. Hopkins (Eds.), *Poissons d'eaux douces et saumâtres de basse Guinée, ouest de l'Afrique centrale/The fresh and brackish water fishes of Lower Guinea, west-central Africa. Vol. 1. Coll. Faune et Flore tropicales 42* (pp. 347–411). Paris, France: Istitut de recherche pour le développement Muséum nationale d'histoire naturelle, Paris, France and Musée royale de l'Afrique centrale, Tervuren, Belgique. p. 800.

Paradis, E. (2010). pegas: An R package for population genetics with an integrated modular approach. *Bioinformatics*, *26*(3), 419–420.

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20*(2), 289–290.

Pereira, L. H., Hanner, R., Foresti, F., & Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse neotropical freshwater fish fauna? *BMC Genetics*, *14*, 20.

Popescu, A. A., Huber, K. T., & Paradis, E. (2012). ape 3.0: New tools for distance based phylogenetics and evolutionary analysis in R. *Bioinformatics*, *28*, 1536–1537.

Posada, D. (2008). jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, *25*, 1253–1256.

Pyron, R. A., & Burbrink, F. T. (2009). Lineage diversification in a widespread species: Roles for niche divergence and conservatism in the common Kingsnake, *Lampropeltis getula*. *Molecular Ecology*, *18*, 3443–3457.

Roberts, T. (1975). Geographical distribution of African freshwater fishes. *Zoological Journal of the Linnean Society*, *685*(57), 249–319.

Rohlf, F. J. (2015). The tps series of software. *Hystrix*, *26*(1), 9–12.

Saitou, N., & Nei, M. (1987). The neighbor-joining method - a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*(4), 406–425.

Schaefer, S. A. (2007). Petersiini. In M. L. J. Stiassny, G. G. Teugels, & C. D. Hopkins (Eds.), *The Fresh and Brackish Water Fishes of Lower Guinea, West-Central Africa* (Vol. I, pp. 380–409). Paris, France, MRAC Tervuren, IRD Éditions Paris: MNHN.

Schliep, K. P. (2010). phangorn: Phylogenetic analysis in R. *Bioinformatics*, *27*(4), 592–593.

Snoeks, J., Harrison, I. J., & Stiassny, M. L. J. (2011). Chapter 3: The status and distribution of freshwater fishes. In D. WRT, K. G. Smith, D. J. Allen, R. A. Holland, I. J. Harrison, & B. EGE (Eds.), *The diversity of life in African freshwaters: Underwater, under threat. An analysis of the status and distribution of freshwater species throughout Mainland Africa* (pp. 42–73). Gland, Switzerland and Cambridge: IUCN.

Sokal, R., & Michener, C. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, *38*, 1409–1438.

Sonet, G., Snoeks, J., Nagy, Z. T., Vreven, E., Boden, G., Breman, F. C., et al. (2018). DNA barcoding fishes from the Congo and the Lower Guinean Provinces: Assembling a reference library for poorly inventoried fauna. *Molecular Ecology Resources*, 1–16. https://doi.org/10.1111/1755-0998.12983.

Song, H., Buhay, J. E., Whiting, M. F., & Crandall, K. A. (2008). Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences*, *105*(36), 13486–13491.

Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, *22*(21), 2688–2690.

Toham, A. K., & Teugels, G. G. (1997). Redescription of two morphologically similar *Brycinus* species (Ostariophysi: Characidae) from Central Africa. *Copeia*, *1997*, 363–372.

Use of Fishes in Research Committee (joint committee of the American Fisheries Society, the American Institute of Fishery Research Biologists and the American Society of Ichthyologists and Herpetologists). (2014). *Guidelines for the use of fishes in research*. Bethesda, MA: American Fisheries Society.

Valdez-Moreno, M., Ivanova, N. V., Elías-Gutiérrez, M., Contreras-Balderas, S., & Hebert, P. D. N. (2009). Probing diversity in freshwater fishes from Mexico and Guatemala with DNA barcodes. *Journal of Fish Biology*, *74*(2), 377–402.

Valentin, A. E., Penin, X., Chanut, J. P., Sévigny, J. M., & Rohlf, F. J. (2008). Arching effect on fish body shape in geometric morphometric studies. *Journal of Fish Biology*, *73*(3), 623–638.

Vilgalys, R. (2003). Taxonomic misidentification in public DNA databases. *New Phytologist*, *160*, 4–5.

van der Walt, K., Mäkinen, T., Swartz, E. R., & Weyl, O. L. F. (2017). DNA barcoding of South Africa's ornamental freshwater fish – are the names reliable? *African Journal of Aquatic Science*, *42*, 155–160.

Ward, R. D. (2009). DNA barcode divergence among species and genera of birds and fishes. *Molecular Ecology Resources*, *9*, 1077–1085.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B*, *360*(1462), 1847–1857.

Witt, J. D., Threloff, D. L., & Hebert, P. D. (2006). DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: Implications for desert spring conservation. *Molecular Ecology*, *15*(10), 3073–3082.

Zanata, A. M., & Vari, R. P. (2005). The family Alestidae (Ostariophysi, Characiformes): A phylogenetic analysis of a trans-Atlantic clade. *Zoological Journal of the Linnean Society*, *145*, 1–144.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Arroyave J, Martinez CM, Stiassny MLJ. DNA barcoding uncovers extensive cryptic diversity in the African long-fin tetra *Bryconalestes longipinnis* (Alestidae: Characiformes). *J Fish Biol*. 2019;1–14. https://doi.org/10.1111/jfb.13987